

Claims

- 5
1. A method of processing distorted short-term speech spectra for automatic speech recognition, comprising:
- 10
- a) providing a set of reference speech spectra;
  - b) determining the reference speech spectra which correspond to the distorted short-term speech spectra;
  - c) estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra; and
  - d) compensating the distorted short-term speech spectra based on the estimated frequency response.
- 15
2. The method according to claim 1, further comprising analyzing the distorted speech spectra by means of a speech/non-speech decision and performing steps b), c) and d) of claim 1 only with respect to the distorted speech spectra which contain speech.
- 20
3. The method according to claim 1, wherein the distorted speech spectra are compensated in the spectral domain or in any domain that can be derived from the spectral domain by a linear transformation.
- 25
4. The method according to claim 1, wherein the set of reference speech spectra is obtained from speech data subject to a known frequency response or subject to low distortion.
- 30
5. The method according to claim 1, wherein the reference speech spectra corresponding to the distorted speech spectra are determined by finding the reference speech spectra closest to the distorted speech spectra.
- 35

6. The method according to claim 1,  
wherein the set of reference speech spectra is constituted  
by model speech spectra from which reference models for  
automatic speech recognition are built.
- 5
7. The method according to claim 6,  
wherein the reference speech spectra corresponding to the  
distorted speech spectra are determined by finding the one  
or more reference models matching a sequence of distorted  
speech spectra and by analyzing which model speech spectra  
matches the distorted speech spectra.
- 10
8. The method according to claim 1,  
wherein for the purpose of determining the reference  
speech spectra corresponding to the distorted speech spec-  
tra the distorted speech spectra are compensated based on  
a previously estimated frequency response.
- 15
9. The method according to claim 1,  
wherein the frequency response is estimated based on the  
difference between the distorted speech spectra and the  
corresponding reference speech spectra.
- 20
10. The method according to claim 9,  
wherein the frequency response is estimated by averaging  
the differences over a plurality of distorted short-term  
speech spectra and the corresponding reference speech  
spectra.
- 25
11. The method according to claim 1,  
wherein a sequence of distorted speech spectra is compen-  
sated based on the frequency response estimated for a pre-  
vious sequence of distorted speech spectra.
- 30
12. The method according to claim 1,  
further comprising smoothing the frequency response esti-  
mated for a sequence of distorted speech spectra taking
- 35

into account the frequency response estimated for a previous sequence of distorted speech spectra.

13. The method according to claim 1,  
further comprising compensating for additive noise in the distorted speech spectra prior to determining the reference speech spectra.
14. A method of processing distorted short-term speech spectra to be used for automatic speech recognition, comprising:
  - a) providing a set of reference speech spectra;
  - b) obtaining distorted speech spectra and analyzing the distorted speech spectra by means of a speech/non-speech decision to filter out the distorted speech spectra that do not contain speech;
  - c) searching the set of reference speech spectra to determine the reference speech spectra which correspond to the distorted speech spectra that have not been filtered out;
  - d) estimating a frequency response by comparing the distorted speech spectra and the corresponding reference speech spectra; and
  - e) compensating the distorted speech spectra using the frequency response estimated in step d).
15. A computer program product comprising program code portion for performing in an automatic speech recognition system the steps of:
  - a) providing a set of reference speech spectra;
  - b) determining the reference speech spectra which correspond to distorted short-term speech spectra;
  - c) estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra; and
  - d) compensating the distorted short-term speech spectra based on the estimated frequency response.

16. The computer program product of claim 15, stored on a computer readable recording medium.

17. A device for processing distorted short-term speech spectra for automatic speech recognition, comprising

- a database for reference speech spectra;
- a processing stage for determining the reference speech spectra corresponding to the distorted short-term speech spectra and for estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra;
- a compensation unit for compensating the distorted short-term speech spectra based on the estimated frequency response.

18. The device according to claim 17, further comprising a buffer for temporarily storing the estimated frequency response.

19. A terminal comprising a speech analyzing stage having

- storing means for storing a set of reference speech spectra;
- processing means for searching the storing means to determine the reference speech spectra that correspond to distorted short-term speech spectra generated within the terminal, and for estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra; and
- compensation means for compensating the distorted short-term speech spectra based on the estimated frequency response.

20. A distributed speech recognition system comprising:
- a) at least one terminal with
    - storing means for storing a set of reference speech spectra;
    - 5 - processing means for searching the storing means to determine the reference speech spectra that correspond to distorted short-term speech spectra generated within the terminal, and for estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra; and
    - 10 - compensation means for compensating the distorted short-term speech spectra based on the estimated frequency response; and
    - 15 b) a network server with central speech recognition means.
- 20